電子化ドキュメントの活用

ノートブック: ICTベーシック／情報リテラシー
作成日: 2015/05/28 23:31 更新日: 2015/05/29 17:37

電子化ドキュメント(PDF)の活用

はじめに

大学におけるICT教育は、知的生産活動にICTを活かせるようになってもらうことを目的としている。知的生産とは、知的成果物（つまりアウトプットされる具体的なモノ）を伴う活動のことで、たとえば難しい本を読むことは、それだけでは知的生産とは言わない。加えて、自分で何かを創り出す必要がある。
電子化ドキュメント(PDF)は、「知的生産の4フェーズ」のうち、「整理・保存」フェーズと、「処理・加工」フェーズにまたがる重要な技術である。これをうまく利用することで、知的生産活動の効率は飛躍的に向上するが、従来、ICTの基礎カリキュラムではあまり取り上げられてこなかった。理由は、ワープロ（文書作成ソフト）、表計算ソフト、プレゼンソフトなどのように、ある特定の目的に特化したテクノロジーでないことと、電子化ドキュメントを操作するアプリ(Adobe Acrobat)の使い方が必ずしもやさしくないことだろう。
ここでは、電子化ドキュメントを、私たちの知的生活に欠かせない2つの情報空間――つまり、電子の世界と紙の世界――を橋渡しし、統合的に扱うためのテクノロジーとして取り扱う。在学中にぜひマスターしてほしい。

電子化ドキュメントの意義

電子の世界と紙の世界をつなぐ

知的生産活動の大半は、ICTが直接扱える「電子の世界」と、従来の紙メディアが存在する「紙の世界」の2つの舞台にまたがって行われる。知的生産のために「収集・入力」する情報が、どちらにも存在するからである。紙メディア上の情報は、時代とともに、電子の世界にコピーされてゆくが（この作業を「電子化」という）、著作権保護や、出版業界の動向や、図書館などの運営を考えると、当分の間、すべての情報が電子化される見込みはないだろう。したがって、私たちはつねに、この2つの世界を相手にしながら知的生産活動をしなくてはならず、そのことがICTの利用法にも大きく影響するのである。
つまり、個人のレベルで電子の世界の情報と紙の世界の情報を一元化する必要がある。
このためには、たとえば、すべての電子データを紙に印刷し、知的生産活動は紙ベースで行うという手段もある。たとえば団塊世代のビジネスマンはそうしていた。紙を満載した「キングファイル」を繰りながら仕事をしている姿がそれだ。
しかしこの科目では、数十年後までの将来を見据えてICTスキルとリテラシーを学ぶのだから、「すべてを紙に」ではなく、逆に「すべてを電子化」し、知的生産活動を電子ベースで行うスキルを身につけたい。そのためのテクノロジーが「電子化ドキュメント」なのだ。
なぜ、このようなものが必要なのか、**図１**で考えてみよう。

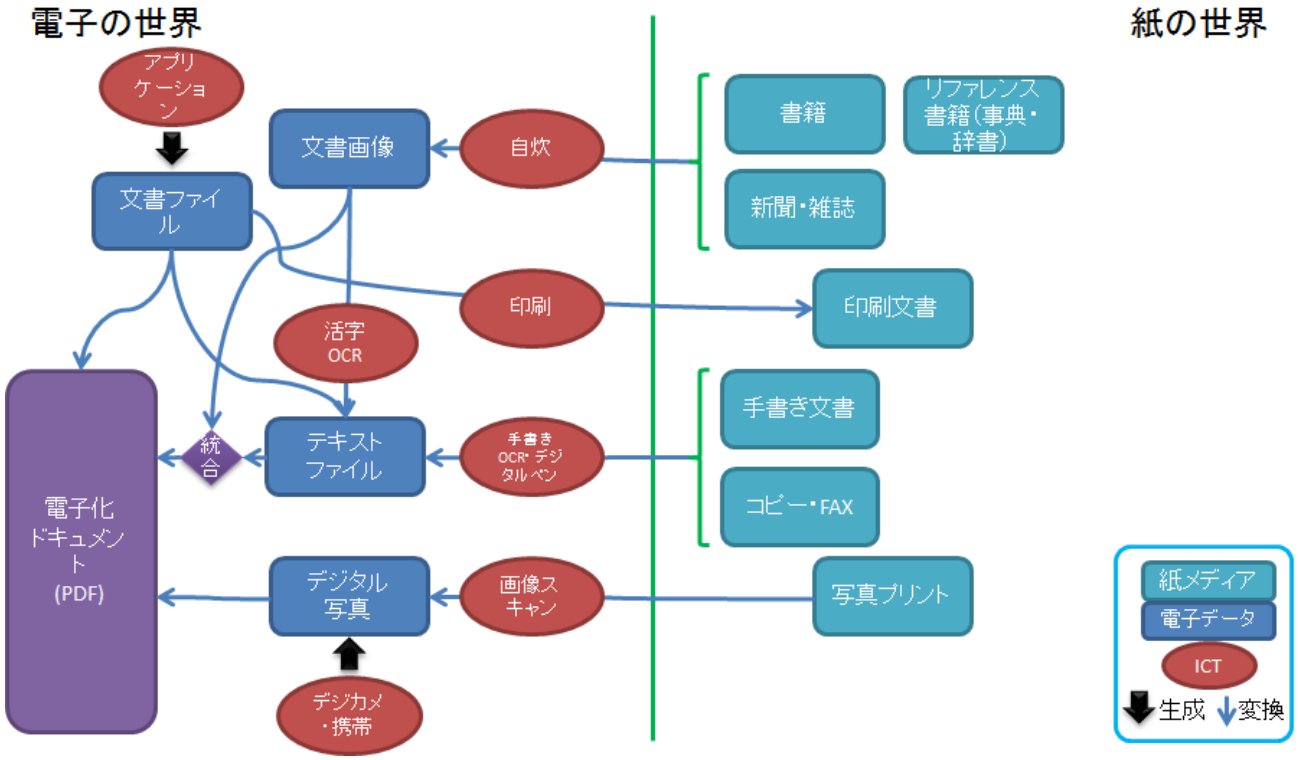

図1：電子化ドキュメントは電子の世界と紙の世界をつなぐ

電子化ドキュメントは、データを紙に印刷できるものなら、どのようなアプリケーションからも作ることができる（PCにAdobe Acrobatがインストールされていなければならないが）。変換したドキュメントは元のアプリケーションとは独立なので、それがインストールされていないPC環境でも読める。この特徴から、電子メールの添付文書や、企業が取扱説明書や申込書といった文書を顧客にインターネット経由で提供するためのフォーマットとして多用されている。また、変換したものは元のアプリケーションでも編集できないが、これは逆に、一度「発行した」文書の内容が改ざんされないという利点につながっている。
また、紙メディア（書類、書籍など）をスキャンして文書画像データにしたものや、デジタル写真データからも変換することができる。文書画像データはいわば「不完全な電子化データ」であるが、後述のAdobe Acrobatを用いて活字OCRにかけると、文書画像データと読取ったテキストファイルのデータが表裏一体となった形式（透明テキスト）で保存できる。これらの電子化ドキュメントは、アプリケーション由来の電子化ドキュメントと何ら区別なく扱えるので、「電子の世界の情報と紙の世界の情報を一元化する」役割を担うことができるのである。

情報（データ）の長期保存と共有

つぎに、知的成果物や中間作成物としての電子データを長期間に亘って保存したり、他人と共有したりすることを考える。
これまでに学んだワープロや表計算といったアプリが作成する「文書ファイル」というデータは、そもそも情報を長期に保存したり、他人と共有するのにすら不適当である。これらのデータは、それを作成したアプリがないPC環境では読むことができないし、バージョンアップやベンダーの倒産といった外的事情によっても旧式化してしまう可能性があるからである。
このことは、現在日本で普及が進まない電子化書籍についても当てはまる。紙

の本であれば、1度購入すれば、出版社がどうなろうと永久に「自分の本」だが、電子化書籍は、買った後でも、出版社や読書アプリのベンダーの事情で、ある日突然読めなくなるかもしれない。「買っても、自分のものになった気がしない」という性質が、日本で電子化書籍の普及が進まない最大の原因だろう。
ならば、電子化データを、長期にわたって安全に保存し、誰とでも共有するにはどうしたらよいか。1つの方法は、文書ファイルをテキストファイルに変換して保存することだが、それでは文書中のマルチメディア要素（画像など）が失われてしまう。
そこで便利なのが、各アプリケーションの文書を、電子化ドキュメント(PDF)に変換して保存しておくことである。この方法なら、紙に印刷したままのイメージで長期間保存できる。

Adobe ReaderとAdobe Acrobat
電子化ドキュメントを扱うには、つぎの2つのアプリケーションを、必要に応じて使い分けなければならない。

- Adobe Reader：電子化ドキュメントを読むためのビューワアプリ。「アプリがなければ読めない」のは、Wordなど通常アプリと同じなのだが、各種プラットフォーム用のものが無料で配布されており、誰でも入手してインストールできるので、通常アプリとは性格が異なる。また、ベンダー（Adobe社）の倒産に際しても、PDF形式は電子化ドキュメントのデファクト・スタンダードとなって久しいため、突然読めなくなる恐れはほとんどなく、万一の場合もドキュメントを他の形式に変換する手段が提供されるのは間違いない。
- Adobe Acrobat：電子化ドキュメントを作成・編集するのに必要なアプリ。有料だが、大学の実習用PCにもインストールされているし、学生のうちは「アカデミック・ライセンス」で安く購入できる。前述のように、電子の世界と紙の世界にまたがる作業が知的生産には必須なので、本格的に取り組むには、在学中に購入することをお奨めする。

電子化ドキュメントを作ってみよう
アプリからの作成
つぎのようなアプリから、電子化ドキュメントを作成してみよう。

- ワープロ(Word)
- 表計算ソフト(Excel)
- プレゼンソフト(Powerpoint)
- ブラウザ(Internet Exproler)

デジタル写真からの作成
携帯やスマホ、デジカメで撮影した画像ファイル（jpeg形式）の画像を何枚か束ねて1つの電子化ドキュメントのファイルを作り、適当な見出しをつけてアルバムにしてみよう。

書籍からの作成
電子化ドキュメントの、有効な利用法の1つは、紙の書籍をまるごと電子化し、1つのファイルに変換することである。この作業を「自炊」という。「自分でデータを吸い上げる」意味だそうで、適切な命名とも思えないが、すでに普及してしまっ

ている。

- 指定するフォルダにある、いくつかの電子化書籍を閲覧してみる。
- 各種機材を用いて、1冊の本を電子化してみる。

自炊はまだ新しい技術であるため、その可能性や問題点を巡って議論が多く、一部では社会問題化しているとさえ言える。そこで、後述の「自炊実習」の節で改めて論じる。

注釈とレビュー

電子化ドキュメントには、マークや注釈を書き込める。これは紙の文書における、ラインマーカーでのマーク付けや文字の書き込み、付箋の貼り付けなどに相当する機能である。電子化ドキュメントであるから、跡形もなく復元することもできる。「書籍からの作成」の項で作った電子化書籍をAcrobatで閲覧し、注釈をつけながら読んでみよう。
また、文書をインターネット上で共有して、複数の人間が同時にレビューすることもできる。これには、

- 電子メールベースのレビュー
- ブラウザベースのレビュー

の2通りがある。どちらのレビューでも、各参加者の環境でPDFファイルに書き込みができること、すなわちAdobe Acrobatがインストールされていることが前提になる。
電子メールベースのレビューは、電子化ドキュメントを添付したメールを参加者同士でやりとりし、文書に注釈を加えることで進められる。これは学会の研究会のようなルーズな集団でのレビューや、文書に多くの改変が予想されない場合に向いている。
ブラウザベースのレビューは、電子化ドキュメントを、特別な機能を持ったWebサーバーか、またはネットワーク上の共有ディレクトリに置いておき、それを参加者がブラウザで閲覧しながら注釈を加えることによって進められる。これはビジネスにおける企画書のレビューのように、文書に大幅な改変が予想され、特定のメンバーで集中的にレビューを進めるのに向いている。

自炊実習：「本」を物質からエネルギーに乗せ換える

電子メディアと「自炊」

「自炊」の語源は、本を電子データに「自分で吸い上げること」らしい。情報学的には、書き言葉という記号を、紙とインク（物質）から磁気記録／光の濃淡（エネルギー）へ「担荷体変換」することと定義できる。
自炊によって本をエネルギーの世界に送り込むことの利点は計り知れない。たとえば、つぎの点が挙げられる。

- 本の保存場所が0になる。
- 本の寿命は永遠になる。
- さまざまなデバイスで読めるようになる。

    - パソコン、スマホ、iPad、ウェアラブルPC……
    - すでに自炊済みの本が、将来登場するデバイスで読める可能性も十分ある。
- 多重分類、内容検索など、ICTが活用できる。

利点が大きい反面、（潜在的な）問題点も大きい。
自炊した電子化ドキュメントは、いわば「自家製の電子化書籍」である。これは、市販されている電子化書籍とは異なり、違法な利用法を防止するセキュリティ上の仕組みがないため、個人で自由に扱える反面、メディア・リテラシー意識をしっかり持っていなければ、つい法律に触れてしまう危険もあるので、十分な注意が必要である。たとえば、次の行為はいずれも違法である。

- 図書館で借りた本を自炊する。
- 自炊した本のファイルを友人に貸したり、ネットにアップロードする。

出版社側は、こうした行為の横行により、以下のような危惧を抱いている。その一部は現実になっている。

- 電子データされた「本」がネット上に流出すると、著作権が侵害される
    - 海外ダウンロードサイト上には膨大な「ジャパン・コミックス」が存在
- 本の寿命が永遠になることで、ビジネス機会が失われる

「自炊」の技法：破壊的自炊

「自炊」という語でこの破壊的自炊を指すこともある。いわば狭義の自炊である。この技法では、電子化された紙の本は裁断されて、単なる紙の束になってしまうので、普通は廃棄される。

- 本のとじしろを裁断機で切り離す
    - 「本」は「紙の束」になってしまう。
- 「ドキュメントスキャナ」で両面の画像を読取り、電子化する
- 本1冊が50～200MB程度の電子化ドキュメント（PDFファイル）に変換される
    - 8TBのハードディスク1台に4万冊～16万冊格納可能
    - 1冊につき3～6分で終了（読み取り速度25枚/分の場合）

図2に、破壊的自炊作業で利用される機器（裁断機、ドキュメントスキャナ）と電子化された本の読書イメージを示す。紙の本と同様に付せんを貼ったり、書き込みができるほか、内容を検索したりもできる。

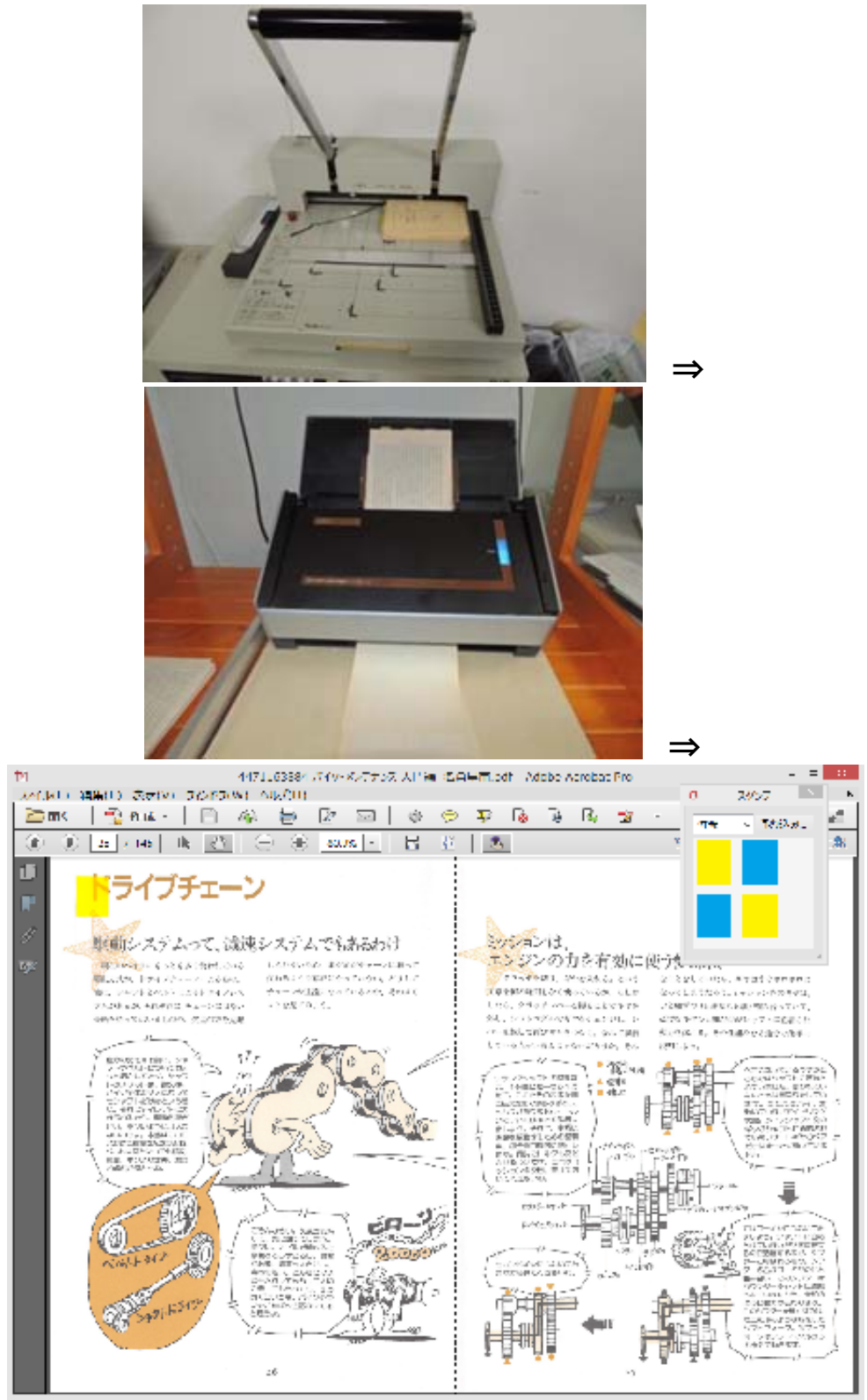

図2：自炊作業で利用される機器と読書画面

「自炊」の技法：非破壊的自炊

非破壊的自炊とは、本を裁断しない。つまり紙の本として残したい場合の技法である。

- 「ブックスキャナ」でページ画像を読取り、電子化する
  - 「のど」の際2mmまでスキャン可能な、特殊なフラットベッドスキャナ

である。
- 変換結果は破壊的自炊と同じ
  - 1冊につき30分で完了。破壊的自炊より能率は悪い
  - 図書館の本などを自炊できてしまう問題点がある（著作権法違反）

図3に、非破壊的自炊作業で利用される機器（ブックスキャナ）を示す。読書イメージは破壊的自炊と同じである。

図3：非破壊的自炊作業で利用される機器

「自炊」をめぐる法的論争

「自炊」は著作権法にいう「私的複製」の範囲内で、法的に問題ない。しかし、自炊設備を持ち、有料で自炊作業を代行する「自炊代行業者」の登場により、自炊が社会問題化した。顧客は電子化書籍を電子メールなどで受け取る一方、業者の手元にも電子化書籍と、裁断済の本が残るからである。もしも業者が裁断済の本を再度自炊すれば、1冊の紙の本から無制限にコピーができてしまうことになる。
そこで、作家団体（日本文藝家協会など）や、出版社は、自炊代行業者を相手取って自炊行為の差し止めを求める訴えを東京地裁に起こした（2011年12月20日）。その主張の骨子は、代行業者による自炊は「私的複製」ではなく、著作権侵害にあたるというものである。
作家側の言い分。

- 「『電子書籍を出さないからスキャンするんだ』という業者にはこう言いたい。『売ってないから盗むんだ』、こんな言い分は　通らない」（東野圭吾）
- 「裁断本を見るのは本当につらい。本当に、もうちょっと本を愛してくださいといいたい」（武論尊）

2012年05月22日　代行業者2社が自主的に業務停止、会社を解散したのを受け、原告弁護団が訴訟を取り下げ、原告側の実質的勝訴で終結した。

「自炊」技術の可能性と危険性

この訴訟の経緯は、自炊技術の相反する両面をを考えるよいきっかけになる。

- 可能性
  - 「自炊」により、個人の読書スタイルは飛躍的に自由度を増す

    - 本というデータ自体にも付加価値が増す
    - 本の「ヘヴィ・ユーザ」にとっての空間的・金銭的恩恵は計り知れない
- 危険性
    - つい法律に触れ、著作権者に害を与える危険性がある
    - 自炊された書籍データが、ネット上に流出することで、被害がさらに増大する

大切なことは、危険性をよく知り、弊害を食い止めながらも、「自炊」技術を単に反社会的なものと見なして、可能性の芽を摘み取らないようにすることではないだろうか。